

# Geographic Information Analysis of Pediatric Lead Poisoning

Florence Lansana Margai, PhD\*

Department of Geography, Binghamton University-SUNY, Binghamton, NY

## Abstract

This study analyzes the spatial distribution of pediatric lead poisoning cases in relation to environmental indicators of lead, housing, and the demographic attributes of block groups in Binghamton, New York. Primary data on childhood blood lead levels are based on screening records from July 1991 to June 1995. Approximately 17% of all children tested within this period had elevated blood lead levels. A number of geographic information system (GIS) and statistical operations are used to determine (a) whether the distribution of lead poisoning cases reflects a consistent spatial pattern, and (b) the extent to which the pattern is linked to possible sources or pathways of exposure such as lead emitting facilities, major transportation corridors, trace lead in soil and municipal water supply, and housing. The results reveal clearly defined clusters of lead poisoning cases along transportation lines within the urbanized and industrialized zones. Specifically, block groups in the central city that were characterized by old, subdivided, and rented properties and poverty had proportionately higher incidences than others. Nearly six out of every ten cases fell within these clusters. These results demonstrate how comprehensive health and environmental data can serve as input in delineating high-risk areas for lead monitoring and remediation programs.

Keywords: pediatric blood lead levels, lead poisoning, GIS statistics, canonical correlation

## Introduction

Lead poisoning is one of the most significant pediatric environmental health hazards in the United States, yet it is one of the most preventable as well. Over the years, several steps have been taken at different fronts to minimize the risks associated with this hazard. The enactment of the Clean Air Act in the 1970s and subsequent federal regulations have led to a more than 90% reduction in atmospheric lead levels. Unfortunately, this metal continues to pose a significant health threat from pre-existing sources such as lead paint and dust in older housing, industrial emissions, and contaminated soils. Several studies conducted over the last two decades have consistently identified neurological and developmental problems in children exposed to lead, even those exposed to levels once considered to be harmless (1,2). Estimates based on the standards set by the Center for Disease Control (CDC) suggest that approximately 10 million children are at risk (3). This is particularly severe in the inner cities, where more than 60% of low income and minority children are believed to have elevated blood lead levels of 10 or more micrograms per deciliter ( $\mu\text{g}/\text{dL}$ ). Explanations for the observed spatial patterns have been linked to historical patterns of urbanization, transportation, and industrialization (4).

\* Florence M. Lansana Margai, Department of Geography, Binghamton University-SUNY, Binghamton, NY 13902 USA; (p) 607-777-6731; (f) 607-777-6456; E-mail: [Lansana@Binghamton.edu](mailto:Lansana@Binghamton.edu)

One of the primary reasons for conducting this study was to explore the use of geographic information systems (GIS) in developing a pediatric lead prevention program for the city of Binghamton, New York. Specifically, the following research questions were examined:

1. What is the spatial distribution of elevated blood lead incidences among children? Is it random or spatially clustered?
2. Is there a significant relationship between the observed distribution of child elevated blood lead incidences and the demographic attributes of residents?
3. Is there a relationship between the observed distribution of child elevated blood lead incidences and significant sources and pathways for environmental lead, such as paint, soil and water quality, transportation corridors, and lead-related businesses and industries?
4. Where are the high-risk areas and what is the density of preschool children within the proximity of these sites?
5. What level of monitoring is necessary to reduce the risks of lead poisoning?

Addressing these questions required the assistance of a GIS and spatial statistics. Using data from different sources, we were able to identify spatial patterns in elevated blood lead incidence cases by block group, inventory the multiple sources of lead contamination, and statistically evaluate the underlying relationships between lead poisoning and various indicators of environmental lead.

### **Developing a GIS for Pediatric Lead Poisoning Prevention**

Recent trends show a growing number of lead monitoring and prevention programs based on guidelines established by the CDC (5,6). Children aged six months through six years are screened regularly at well-child visits. Families of children with elevated lead levels are given prompt medical attention coupled with residential testing to identify and possibly eliminate the source of lead. While these efforts are successful in curbing the rate of lead poisoning, there are still some problems. For example, even though the CDC recommends universal screening, not all children are being tested. Even among those who are tested, no effort is made to educate the parents about the dangers of lead poisoning unless the test results are positive. As Wartenberg (7) argues, such an approach not only hinders the primary prevention efforts but the regional assessment of risks as well. Neighbors may not be readily identified and geographic clusters of highly exposed individuals are likely to be missed.

Children are exposed to lead from multiple sources. Therefore, a major step toward the development of an efficient lead monitoring and prevention plan must start with a comprehensive database that includes not only the screening records of children but ancillary data on industrial emissions, transportation lines, housing characteristics, occupational exposure patterns, and other parameters. The use of a GIS can facilitate this process. A GIS is essentially a collection of computer hardware and software that can be used to capture, store, retrieve, analyze, and visualize various forms of spatial data. It is a very powerful tool for understanding the spatial linkages between multiple layers of human and natural phenomena and for isolating possible cause-and-effect relationships. On matters relating to public health, it can serve as a modeling tool for spatial epidemiological patterns as well

as an analytical tool for testing hypotheses regarding mapped distributions of disease (8).

The methodology for compiling various sources of lead toxicity into a GIS is still in its infancy, however. Few researchers have fully explored this technology as a reliable means of identifying lead exposure patterns and high-risk areas (4,7,9,10). Some of the applications have been exploratory, with limited data used to map exposure patterns at coarse and sometimes inappropriate spatial levels such as zip codes, minor civil divisions (MCDs), or census tracts. It is important, however, to go beyond this exploratory stage and fully utilize the analytical and predictive functions of GIS in discerning potential risk areas.

### The Study Area

Binghamton extends across 11 square miles with approximately 53,000 people. This city grew out of an extensive industrial heritage. Some of the businesses with roots in this area include International Business Machines (IBM), Endicott Johnson Shoe Corp., General Electric, Universal Instruments, Ozalid, Link Federal Systems, and Anitec. The rise of these companies brought an industrial boom that required extensive road and rail systems for transportation as well as mass inexpensive housing. Today, Binghamton is criss-crossed by the Canadian-Pacific railway system, the railroad network in northeastern United States, and at least four major highways: US 11, NY 17, Interstate 81, and Interstate 88.

Binghamton is the most urbanized area in the two-county southern tier of New York State. About 88% of the population is white, with minorities constituting the remaining 12%. Like several other northern US cities, the city has experienced population and economic declines. Several industries have closed or downsized their workforce to cope with economic difficulties. There has also been a drop in downtown activities and service functions due to suburbanization and the building of shopping centers and strip malls in the outlying areas. The city today reflects the characteristic patterns of urban decay. Many of the old buildings and homes have been renovated into multiple housing units to serve low-income residents and college students. Within the two-county area, most of the government subsidized housing units are located in the city. Higher income housing and households are located in the outskirts of the city, particularly in the far west and southwestern sides.

### Data Collection

The primary data on pediatric blood lead levels consisted of 1,840 records of children tested between 1991 and 1995. The screening tests were typically for children between one and three years, although a small percentage (6%) of the affected children were five or six years old. Previous studies have indicated that lead is prevalent among children in these age groups primarily because of the increased hand to mouth activities.

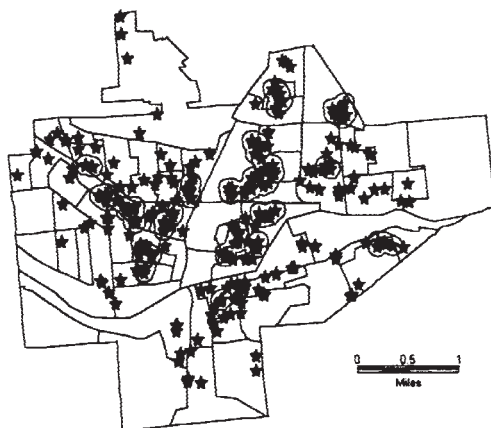
Initial assessment of the data suggested that at least 17% of the children had high blood levels. Based on the CDC guidelines, almost two-thirds of the children were classified at Level IIa, with a blood lead level concentration of 10–14.9 µg/dL. About 24% of the children are classified as Level IIb, with blood lead levels of 15–19 µg/dL. Another 16% of the cases are categorized as Level III, with blood lead levels of 20–44.9

$\mu\text{g}/\text{dL}$ . Fewer than 1% of the children are in Level VI, the highest category observed in the community, with blood lead levels of 45–69  $\mu\text{g}/\text{dL}$ . The majority of the children are white, reflecting the racial composition of the population in the city. Based on the addresses, the data were geocoded and exported into a desktop mapping package, MAPINFO, for cartographic analysis.

## Data Analysis and Results

### *The Spatial Distribution of Lead Poisoning Cases*

The spatial distribution of lead cases was evaluated by two methods. First, a buffer analysis was performed using the MAPINFO software. The criterion used for detecting spatial clustering was proximity. A lead cluster was inferred in every area where there were four or more confirmed cases of lead poisoning within a 500 foot radius. The results were then validated using cluster analysis within the statistical software, SPSS. Figure 1a shows the location of these clusters at the block group level. About 57% of the lead poisoning cases fell within 13 defined clusters. Most of the clusters were located close to the center of the city and along the transportation lines. The relationships between these lead clusters and the potential sources of exposure were subsequently assessed using a number of statistical procedures.



**Figure 1a** High blood lead occurrences.

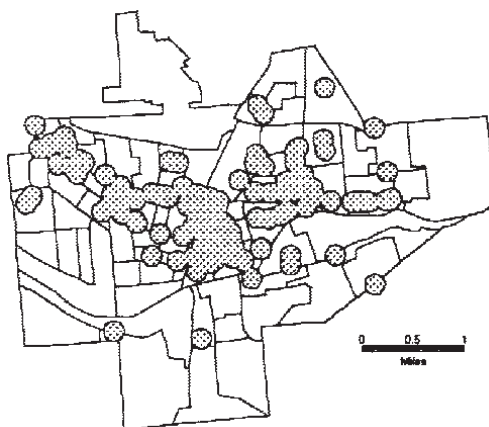
### *Housing and Socioeconomic Correlates of Lead Poisoning*

The source of lead largely depends on the characteristics of the community in which a child resides. Previous studies have reported high lead levels in areas of physical deterioration with old, subdivided housing, poverty, and areas of minority concentration. In an attempt to identify such areas in Binghamton, 12 demographic variables were selected using US Census data from STF3A files. A correlation analysis was then performed between the variables and the rate of blood lead incidences within each block group. Among the 12 variables, 9 were significantly related to pediatric lead poisoning cases in Binghamton. The strongest indicators were poverty and housing quality

variables such as block groups with pre-1940s housing, rented property, and subdivided units. Initially, no associations were observed between the lead cases and areas with significant minority and family composition. However, further analysis using only the lead cases that fell within defined clusters suggested a possible link with the proportion of African Americans in the community ( $r=0.41$ ;  $p<0.01$ ).

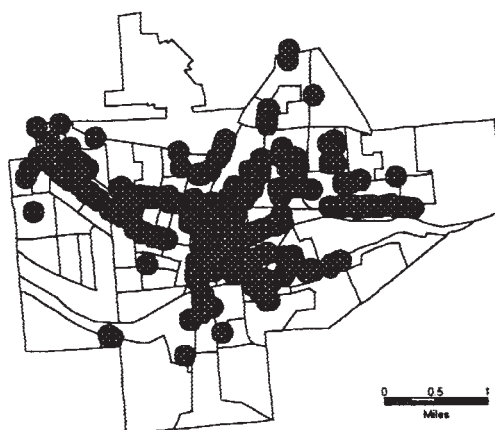
### ***Environmental Sources of Lead Poisoning***

Different sources of data characterizing lead-emitting businesses and industries in the city, automobile-related facilities, and transportation corridors were incorporated into the GIS. First, historical data consisting of all business locations within Binghamton since 1890 were queried. Businesses that qualified for entry into the analysis included those that used lead or lead by-products in their activities. These included factories such as machine shops, foundries, and parts and glass manufacturers. Using the location of these facilities, a 500 foot buffer was established as a reasonable distance over which the airborne effects of lead would be dispersed on land. These buffers covered about 20% of the city's areal extent and about 41% of the confirmed cases fell within the defined buffer (Figure 1b). Further spatial analysis involved subdividing the buffer into smaller polygons with boundaries corresponding to the block group boundaries. The area of each buffered portion was then divided by the total area of the corresponding block group to determine the degree to which each block group was characterized by lead-associated industries or businesses. This newly created variable showed a highly significant relationship with child lead poisoning rates ( $r=0.61$ ;  $p<0.05$ ).

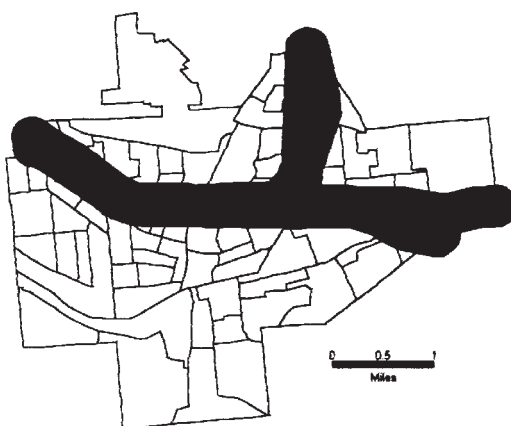


**Figure 1b** Lead emitting/handling businesses.

Using procedures similar to those explained above, a buffer was established for automobile-related facilities such as gas stations, repair shops, dealerships, and junkyards. About 53% of all confirmed cases fell within these areas (Figure 1c). The relationship between the buffered variable and lead poisoning cases was also significant ( $r=0.5$ ;  $p<0.05$ ). Buffers were also created around major roads and railways in Binghamton (Figure 1d). The lead cases were significantly related to both the road buffers ( $r=0.30$ ;  $p<0.05$ ) and the railroad buffers ( $r=0.46$ ;  $p<0.05$ ).



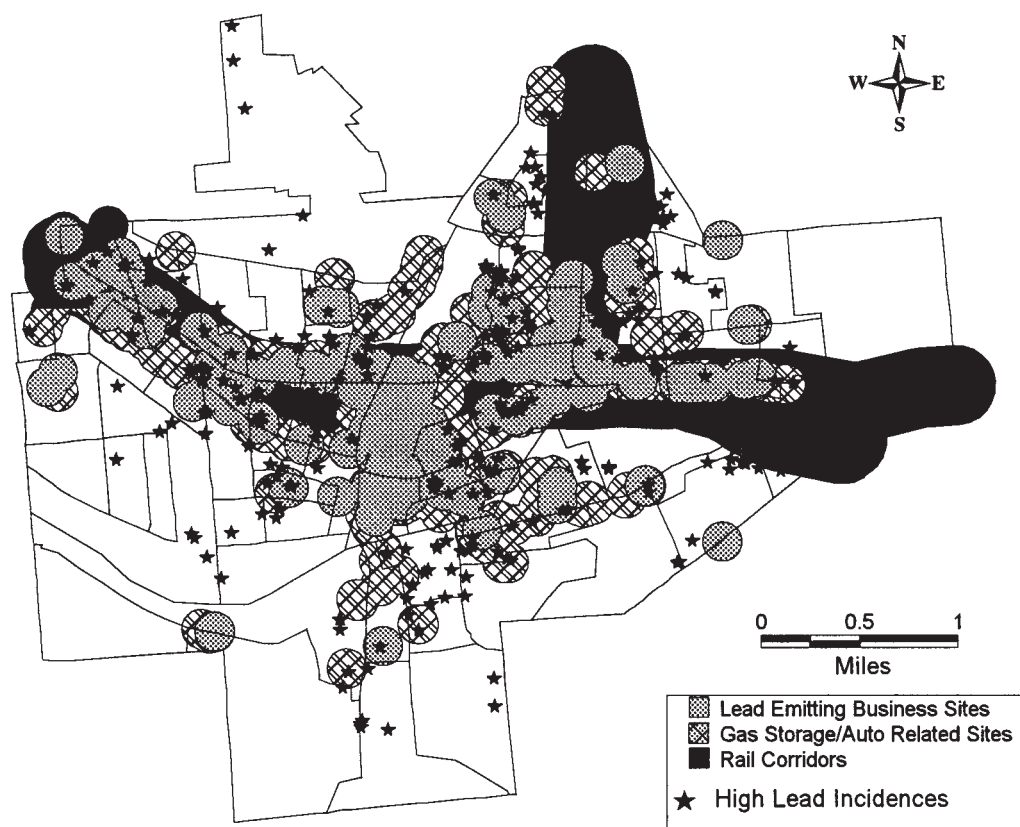
**Figure 1c** Gas storage/auto-related sites.



**Figure 1d** Rail corridors.

### ***Environmental Pathways***

The visual information provided in Figure 1e reflects the spatial variability of pediatric lead poisoning relative to the industrial locations, automobile-related sites, and rail corridors. However, the use of a GIS in lead monitoring and prevention must include not only the potential exposure sources illustrated in the map, but also the pathways that are likely to directly affect the children. Specifically, contaminated soil and water are major pathways for young children. In Binghamton, the threat of trace lead in the water supply had been minimized by the preventive steps taken earlier by the city to minimize the corrosion from pipes. Certain polyphosphate compounds (with commercial brand names such as Aquamag and Calciquest) that bond to water pipes were first added to the municipal water supply in 1992. The impact of these additives was assessed by examining lead poisoning incidences before and after the changes were implemented. While the total number of reported cases varied from year to year, the mean blood lead levels in children declined consistently over time from 15.86  $\mu\text{g}/\text{dL}$  in 1991/1992 to 13.72  $\mu\text{g}/\text{dL}$  in 1994/1995.



**Figure 1e** High lead incidences in all three impact zones.

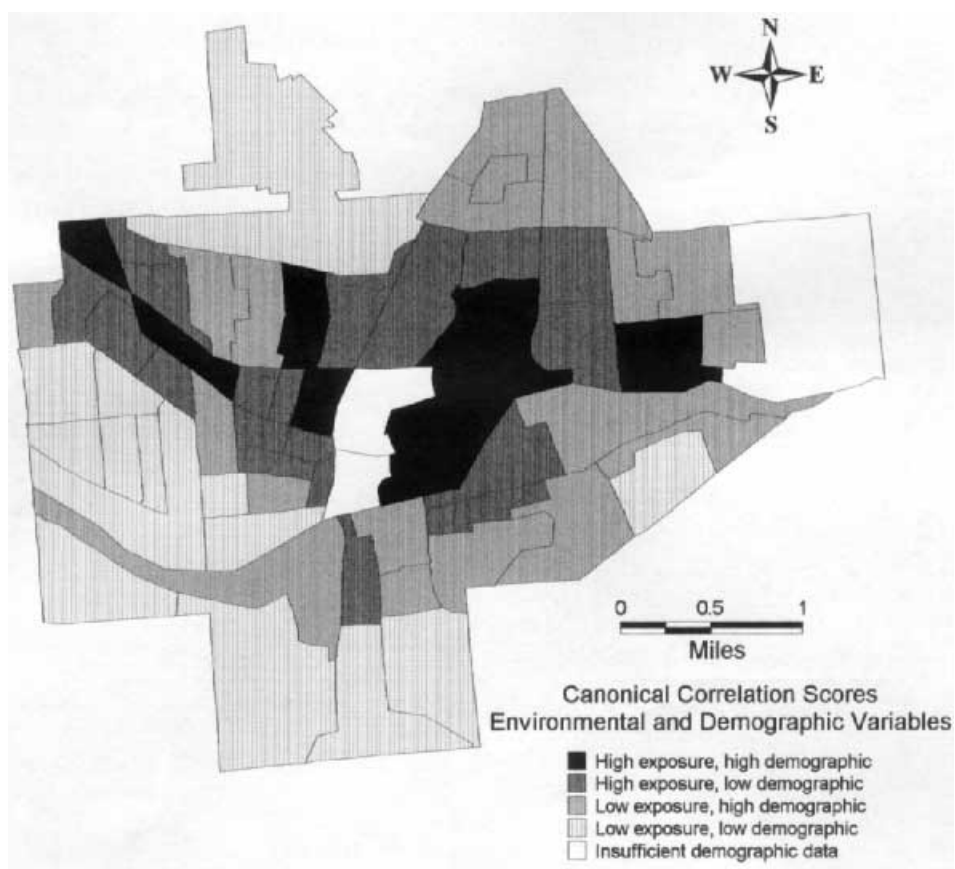
The effect of soil lead on childhood blood lead levels was also evaluated by conducting a detailed soil sampling analysis in July 1996. Soil samples were extracted from the front, rear, and side yards of homes within the lead clusters identified earlier in the study, as well as outside of the clusters. The latter served as the control group. The soil samples were tested in a professional laboratory using the EPA 6010 Method. The results were later integrated with the existing data in the GIS. A query was then performed to identify all sites that exceeded the EPA standards of 500 parts per million. Several of the sites were above the EPA standards for lead in the soil. Lead levels within the clusters found in the central city were two or more times higher than the EPA standards. Statistical analysis of the mean differences confirmed that soil lead levels were significantly higher in the clustered areas than outside of them ( $t=3.66$ ;  $p<0.05$ ). However, the relationship between soil lead and childhood blood lead levels was not significant.

#### ***Delineation of Target Communities for Lead Prevention***

The final objective in this study was to delineate the high-risk areas based on the comprehensive database described above. This was accomplished by using canonical correlation analysis to quantify the associations between the major lead indicators and

then develop scores that would provide the most explanation for the spatial occurrence of the observed lead clusters. Variable selection for this phase was based on statistical significance from the preceding analyses. Two sets of variables were used. The first set consisted of the three variables that measured the effects of railroads, businesses, and automobile-related sites. The second set included the six demographic variables that were best associated with the location of lead poisoning cases. These variables were entered into the canonical correlation procedure.

At the conclusion of the statistical analysis, the canonical coefficient that maximized the linear relationship between the two sets of variables was selected. This coefficient was very high ( $r=0.83$ ), implying that at least 80% of the observed lead cases could be jointly explained by these variables. Pairs of canonical correlation scores, representing the aggregate values of the two sets of variables, were also obtained for each block group. Scores with values greater than zero were classified as HIGH and those with values less than zero were classified as LOW. The results were then mapped to visualize the spatial relationships (Figure 2). As expected, block groups with high scores were found mainly in the center of the city and along the rail corridor going westward. Those



**Figure 2** High-risk areas for lead poisoning in Binghamton, NY; canonical correlation results by block groups.

with low scores on both groups of variables were along the outskirts of the city, almost forming a continuous ring. A few block groups showed dissonant canonical correlation scores (high/low or low/high), but overall, the results showed that among the 26 block groups with high canonical scores on both variables, 191 cases of lead poisoning were reported—a rate of 7.3 occurrences per block group. Among the 29 with low scores on both sets of variables, about 85 cases were found with a rate of 2.9 cases per block group. These results confirm that lead poisoning cases are closely grouped in space and not merely random occurrences. Furthermore, these findings demonstrate the strength of two sets of variables in identifying high-risk areas.

## Conclusions

In building on the strengths of previous applications, this study has examined the spatial patterns of lead poisoning and, within the context of environmental and demographic variables, isolated the high-risk areas for lead intervention programs. All of these significant steps were made possible through the use of GIS and statistical analysis. Several steps were involved, starting from data collection, storage, analysis, and visualization. Obviously one bottleneck in the process was the constant shift of data back and forth from one software to the other. This is not unusual, however, and future developments in the GIS packages are likely to provide improvements in these statistical procedures. Aside from these minor impediments, the technology provides a good basis for handling spatial epidemiological problems.

## Acknowledgments

The contributors to this research were John W Frazier, Steve Walter, and Ron Brink. Assistance with soil sampling analysis was provided by the McNair Scholars Program at Binghamton University. An earlier and more detailed discussion of the study was published in the *Applied Geographic Studies Journal*, Vol.1, No. 4, pp. 253–270.

## References

1. Mushak P. 1992. Defining lead as premiere environmental health issue for children in America: Criteria and their quantitative application. *Environmental Research* 59:281–309.
2. Sciarillo WG, Alexander G, Farrell KP. 1992. Lead exposure and child behavior. *American Journal of Public Health* 82:1356–60.
3. Sargent JD, Braun MJ, Freeman JL, Bailey A, Goldman D, Freeman DH. 1995. Childhood lead poisoning in Massachusetts communities: Its association with sociodemographic and housing characteristics. *American Journal of Public Health* 85:528–34.
4. Bailey AJ, Sargent JD, Goodman DC, Freeman J, Brown MJ. 1994. Poisoned landscapes: The epidemiology of environmental lead exposure in Massachusetts children, 1990–1991. *Social Science Medicine* 39:757–66.
5. Agency for Toxic Substances and Disease Registry (ATSDR). 1988. *The nature and extent of lead poisoning in children in the United States: A report to Congress*. Atlanta: US Department of Health and Human Services. Pub. 99–2966.
6. Centers for Disease Control. 1991. *Preventing lead poisoning in young children: A statement by the Center for Disease Control*. Atlanta: US Department of Health and Human Services. October.

7. Wartenberg D. 1992. Screening for lead exposure using Geographic Information System. *Environmental Research* 59:310–17.
8. Cliff AD, Haggett P. 1996. The impact of GIS on epidemiological mapping and modeling. In: *Spatial analysis: Modeling in a GIS environment*. Ed. by P Longley, M Batty. Cambridge, UK: GeoInformation International. 321–43.
9. Guthe WG, Tucker RK, Murphy EA, England R, Stevenson E., Luckhardt JC. 1992. *Environmental Research* 59:318–25.
10. Griffith DA, Doyle PG, Wheeler DC, Johnson JL. 1998. A tale of two swaths: Urban childhood blood-lead levels across Syracuse, New York. *Annals of the Association of American Geographers* 88(4):640–65.